UNITED STATES PROVISIONAL PATENT APPLICATION

TITLE: Method and Apparatus for Functional Architecture of Voice-over-IP SIP Network Border Element

INVENTORS: Radhika R. Roy, Siroos K. Afshar, Cameron Scott Blandford, Behzad Nadji, Alireza Faryar, and Burwell Goode

[0001] The present invention relates to implementations of signaling protocols such as the Session Initiation Protocol (SIP). See J. Rosenberg, H. Schulzrinne et al., "SIP: Session Initiation Protocol," Internet Engineering Task Force (IETF) Network Working Group, Request for Comments 3261 (June 2002), which is incorporated by reference herein.

[0002] Constructing a large-scale Voice-over-Internet-Protocol (VoIP) network requires careful consideration of economies-of-scale. If a service provider wants to offer services building intelligent services infrastructure efficiently with economies-of-scale over a large network that, for example, spans five continents serving over billions of customers, using a protocol such as the peer-to-peer SIP protocol, the architectural design of the SIP border element is a critical piece of the overall VoIP Infrastructure.

[0003] The present invention is directed to an advantageous architectural design for a SIP border element of a VoIP network. Functional architectural criteria are disclosed that facilitate the offering of intelligent services by the SIP-based entities of the

service complex transparently to the end users. A SIP border element can be designed in an integrated fashion realizing all the functions in a single physical box without exposing its inner functional entities. This kind of design, however, may not always be scalable in many situations for designing a large-scale network like AT&T's VoIP network. The present invention discloses a variety of advantageous ways of decomposing a SIP border element so as to make it scalable. A SIP border element can be decomposed into a plurality of physical entities for optimization of the large-scale design while they act as a single integrated functional entity logically to execute a set of functions at the border of the VoIP infrastructure.

SUMMARY OF DRAWINGS

[0004] FIG. 1 is a functional architecture of a SIP BE realized as a single physical entity.

[0005] FIG. 2 illustrates the separation of a SIP BE in two separate physical entities: signaling and media.

[0006] FIG. 3 illustrates a single SIP BE signaling entity controlling multiple media entities.

[0007] FIG. 4 illustrates decomposition of a SIP BE media entity into a RTP media entity and an FW/NAT entity.

[0008] FIG. 5 illustrates decomposition of a RTP media entity: into a RTP media entity excluding transcoding function and media transcoding entity.

DETAILED DESCRIPTION

[0009] FIG. 1 is a diagram showing the different functional components of an illustrative VoIP architecture. The VoIP architecture advantageously provides a single common infrastructure for facilitating the development of real-time services with the highest quality and availability, the shortest possible time-to-market, and the lowest cost of operations and maintenance feasible. As depicted in FIG. 1, a Border Entity (BE) resides in the boundary of the VoIP/Multimedia Infrastructure to facilitate all services that to be offered by different functional entities to the end users transparently.

[0010] The BE should perform a variety of functions:

[0011] a. Signaling protocol handling from the end systems/devices/users intelligently and communication with the services entities within the VoIP Infrastructure (e.g., SIP, RADIUS/DIAMETER, HTTP/KPML, etc.);

[0012] b. Acting as the boundary of trust so that no security attack can happen to any entities residing within the VoIP Infrastructure (e.g., SIP-aware Firewall/NAT, Authentication, Call Admission Access Control);

[0013] c. VoIP/Multimedia media handling, detection of any services intelligence that may reside in media streams, and notification of services intelligence to the services entity appropriately for offering services to the end users (e.g., RTP media streams handling, DTMF Detection in Media Streams, DTMF digits notification to the application server);

[0014] d. Media transcoding (e.g., between G-series audio codecs, between H-series Video Codecs).

[0015] It is advantageous to provide a variety of architectural design criteria for the SIP BE where SIP will be used by the end system/device/user for

communications with the SIP BE. SIP and other protocols will be used between the SIP BE and the functional entities within the VoIP Infrastructure as appropriate.

[0016] FIG. 1 shows the functional architecture of the SIP BE that is communicating with various entities outside and inside of the VoIP Infrastructure. There can be many BEs to form the boundary of the VoIP Infrastructure for connecting all the end systems for proving services transparently by the services functional entities residing in the services complex. In this context, it can also be assumed that all functions of the SIP BE stated in items a, b, c, and d are realized in a single physical entity. However, SIP BE will communicate with the end system (e.g., IP PBXs, IP GWs, IP devices, SIP Phones) and the call control element (CCE) using SIP protocol. It may also communicate with the CAC server (e.g., Policy, Security, Accounting) using AAA protocol (e.g., RADIUS/DIAMETER) and policy protocol (e.g., COPS). It may be noted that the SIP BE may transfer services intelligence detected in media streams to the application server (AS) using SIP, HTTP/KPML, and/or other protocols per direction of the CCE.

[0017] However, this realization of all functions in a single physical entities may not always be scalable. For example, RTP media streams may require 100s Gigabits per second throughput bandwidth pipes while the signaling entity requires a good amount of CPU processing capacity. These two categories of requirements suggest that is advantageous to split the BE into two different physical entities: a signaling entity and an RTP media entity.

[0018] FIG. 2 shows how the SIP BE can be realized in two separate physical entities, and the functions of those two entities can be as follows:

[0019] *Functions of Signaling Entity*: It will appear as a SIP proxy to the end systems (e.g., SIP Phones, IP-PBXs, IP-GWs) along with SIP back-to-back-user-agent (B2BUA) capabilities. It will appear as a SIP user agent (UA) to the CCE, media server (MS), or other entities within the VoIP Infrastructure that use SIP protocol.

[0020] It will also act as the SIP-aware firewall (FW)/network address translator (NAT) control proxy for opening and closing pinholes (IP addresses, ports) for SIP signaling and RTP media traffic. However, SIP-aware FW/NAT will reside inside the media entity, and a FW/NAT control protocol (e.g., protocol developed IETF's MIDCOM WG, or others) needs to be used between the signaling entity and the media entity.

[0021] It will also be able to communicate with the CAC policy, security, and accounting server using the appropriate protocols (e.g., COPS, RADIUS/DIAMETER).

[0022] It will also be able to transfer DTMF digits to the AS using SIP, HTTP/KPML, or other protocols as appropriate when these digits are sent by the media entity to it.

[0023] There is also a need for a communications protocol between the signaling and media entity for communications especially for controlling (e.g., when to detect DTMF digits and sending those digits, addresses & ports to which media to be received and routed, how to overcome the startup and failure conditions of media and signaling entity, QOS/performance parameters that need to be maintained in the RTP media streams, etc.) of the media entity by the signaling entity. This controlling protocol used between these two entities can be MGCP, MEGACO, IPDC, or other protocols.

[0024] The signaling entity will also instruct the media entity for media transcoding based on the VoIP QOS policy as dictated by the CCE and/or uploaded from the CAC QOS server.

[0025] *Functions of Media Entity:* It will handle RTP media streams as per direction of the signaling entity. For example, when a call comes, the signaling entity will decide which media-type to be handled, which IP address and port number each media stream to be received and to be routed. For this, a communications protocol (e.g., MGCP, MEGACO, IPDC, or other protocols), as stated in the case of signaling entity,. needs to be used between the media and the signaling entity.

[0026] The media entity will also have SIP-aware FW/NAT that will be controlled by the signaling entity as stated earlier for opening and closing pinholes for signaling and media traffic through it. It will filter all RTP media streams based on the security policies dictated by the signaling entity.

[0027] It will relay the RTP media streams to detect whether any services intelligence like DTMF is there. If it detect services intelligence (e.g., DTMF) in the media streams, it will notify the signaling entity using the communications protocol described earlier as dictated by the signaling entity.

[0028] The media transcoding will also be done in the media entity as instructed by the media entity using the communications protocol as described earlier.

[0029]

[0030] It is interesting to note that, unlike the media entity, a signaling entity can control one or more media entities as the signaling entity will not need extensive media handling throughput capacity, other than more CPU capacity. FIG. 3

shows how a single media entity can control many media entities. A SIP BE signaling entity may control many media entities located 100s of miles away. This decomposition of the SIP BE makes the SIP BE architecture further scalable should a large-scale network need to build up.

[0031]

[0032] The media entity can further be decomposed through separation of the firewall/NAT and RTP media entity where the FW/NAT will handle only the IP traffic while the RTP media entity will handle the RTP media streams. FIG. 4 shows the further decomposition of the SIP BE media entity: into a RTP Media Entity and a FW/NAT Entity.

[0033] The RTP media entity also includes media transcoding and DTMF detection. This decomposition will provide the opportunity to further optimize the media functions. For example, RTP media entity will only be optimized to perform RTP related media relay including transcodings. The FW/NAT will only be optimized to deal with IP packet streams without being aware of the RTP packets.

[0034] Again, a single SIP BE signaling entity can also control many RTP media entities and many FW/NAT entities. For simplicity, those configurations have not been shown. This distribution of the media functions will further help to optimize the design of the SIP BE for designing the large-scale network.

[0035]

[0036] Media transcoding requires a lot of DSP power, and it may be worthwhile to share the media transcoding capabilities among multiple entities. FIG. 5

shows a configuration where the RTP media entity is further decomposed making the media transcoding entity a separate entity.

[0037] However, there needs to be a protocol between the media transcoding entity and the RTP media entity (excluding transcoding) for communications how the transcoding needs to be done. RTP entity will control the transcoding entity while the RTP entity will be controlled by the SIP BE signaling entity.

[0038]

[0039] It is important to note that a single SIP BE signaling entity can control many RTP entities (excluding transcoding), transcoding entities, and FWs/NATs entities. Similarly, a single RTP entity can control many transcoding entities. This decomposition of the SIP BE media functionalities can be used for further optimization of the design of the VoIP network that may span five continents serving over billions of customers like AT&T's network. Again, for simplicity, this configuration has not been shown here.

[0040]

[0041]